

ERGODIC RECURRENCE AND BOUNDED GAPS BETWEEN PRIMES

HAO PAN

ABSTRACT. Let $(\mathcal{X}, \mathcal{B}_\mathcal{X}, \mu, T)$ be a measure-preserving probability system with T invertible. Suppose that $A \in \mathcal{B}_\mathcal{X}$ with $\mu(A) > 0$ and $\epsilon > 0$. For any $m \geq 1$, there exist infinitely many primes p_0, p_1, \dots, p_m with $p_0 < \dots < p_m$ such that

$$\mu(A \cap T^{-(p_i-1)}A) \geq \mu(A)^2 - \epsilon$$

for each $0 \leq i \leq m$ and

$$p_m - p_0 < C_m,$$

where $C_m > 0$ is a constant only depending on m , A and ϵ .

1. INTRODUCTION

Suppose that $(\mathcal{X}, \mathcal{B}_\mathcal{X}, \mu)$ is a probability measure space. Let $T : \mathcal{X} \rightarrow \mathcal{X}$ be a measure-preserving transformation on \mathcal{X} , i.e., $\mu(A) = \mu(T^{-1}A)$ for each $A \in \mathcal{B}_\mathcal{X}$. The classical Poincaré recurrence theorem (cf. [7, Theorem 2.11]) says that for any $A \in \mathcal{B}_\mathcal{X}$ with $\mu(A) > 0$, there exist infinitely many integers $n \geq 1$ such that $\mu(A \cap T^{-n}A) > 0$.

Assume that T is an invertible transformation. In 1934, Khintchine [15] considered the set of recurrence

$$S_{A,\epsilon} = \{n \in \mathbb{N} : \mu(A \cap T^{-n}A) \geq \mu(A)^2 - \epsilon\}.$$

Khintchine proved that for any $A \in \mathcal{B}_\mathcal{X}$ with $\mu(A) > 0$ and any $\epsilon > 0$, $S_{A,\epsilon}$ has a bounded gap, i.e., there exists a constant $L_0 > 0$ (only depending on A and ϵ) such that

$$S_{A,\epsilon} \cap [x, x + L_0] \neq \emptyset$$

for any $x \geq 1$. Notice that the bounded gaps can be arbitrarily large for the varied measure-preserving systems. For example, for any positive integer m , consider the cyclic group $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$ with the discrete probability measure. Let $T : x \mapsto x+1$ and $A = \{0\} \subseteq \mathbb{Z}_m$. Clearly $A \cap T^{-n}A \neq \emptyset$ if and only if n is a multiple of m . So the bounded gap of $S_{A,\epsilon}$ is always m for any $0 < \epsilon < m^{-2}$. For the further extensions of Khintchine's theorem, the readers may refer to [2, 8, 4, 5].

On the other hand, a recent breakthrough on number theory is about the bounded gaps between consecutive primes. In [24], with help the Goldston-Pintz-Yıldırım [10] sieve method and a variant of the Bombieri-Vinogradov theorem,

Key words and phrases. ergodic recurrence; bounded gaps between primes; Kronecker factor.

Zhang showed for the first time that the gap between consecutive two primes can be bounded by a constant infinitely often. In fact, Zhang proved that

$$\liminf_{\substack{p, q \rightarrow \infty \\ p < q \text{ are primes}}} (q - p) \leq 7 \times 10^7.$$

Subsequently, using a multi-dimensional sieve method, Maynard [18] greatly improved the bound 7×10^7 to 600. Nowadays, the best bound is 246 [20]. In fact, with help of the multi-dimensional sieve method, Maynard and Tao independently showed that for any $m \geq 1$, the gaps between consecutive m primes also can be bounded by a constant infinitely many times. That is, there exist infinitely many primes p_0, p_1, \dots, p_m with $p_0 < \dots < p_m$ such that

$$p_m - p_0 < C_m,$$

where $C_m > 0$ is a constant only depending on m . Subsequently, the Maynard-Tao theorem was extended to the primes of some special types [19, 6, 16]. There is a nice survey on Zhang's theorem and Maynard-Tao's theorem written by Granville [11]

Notice that Khintchine's theorem, Zhang's theorem and Maynard-Tao's theorem are all concerning the bounded gaps. It is natural to ask whether we can establish a connection between those theorems. The purpose of this paper is to give a Khintchine type extension to the Maynard-Tao theorem.

Let us consider those n in $S_{A, \epsilon}$ which are shifted primes, i.e., a prime minus one. For $A \in \mathcal{B}_{\mathcal{X}}$ with $\mu(A) > 0$ and $\epsilon > 0$, let

$$\Lambda_{A, \epsilon} = \{p \text{ prime} : \mu(A \cap T^{-(p-1)}A) \geq \mu(A)^2 - \epsilon\}.$$

Theorem 1.1. *Let $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, T)$ be a measure-preserving probability system with T invertible. Suppose that $A \in \mathcal{B}_{\mathcal{X}}$ with $\mu(A) > 0$ and $\epsilon > 0$. For any $m \geq 1$, there exist infinitely many primes*

$$p_0, p_1, \dots, p_m \in \Lambda_{A, \epsilon}$$

with $p_0 < \dots < p_m$ such that

$$p_m - p_0 < C_m,$$

where $C_m > 0$ is a constant only depending on m , A and ϵ .

In particular, we may find infinitely many pairs of primes p, q with $p < q$ such that

$$\mu(A \cap T^{-(p-1)}A), \mu(A \cap T^{-(q-1)}A) > 0$$

and the gap $q - p$ is bounded by a constant C .

Let us see a combinatorial consequence of Theorem 1.1. For any $E \subseteq \mathbb{N}$, define the upper Banach density of E

$$\overline{d}_B(E) := \limsup_{\substack{N-M \rightarrow +\infty \\ N \geq M \geq 0}} \frac{|E \cap [M, N]|}{N - M + 1}.$$

In [21], Sarközy proved that if $\bar{d}_B(E) > 0$, there exist infinitely many primes p such that

$$p - 1 \in E - E,$$

where $E - E = \{x - y : x, y \in E\}$. Further, Bergelson and Lesigne [3] showed $\{p - 1 : p \text{ is prime}\}$ is an enhanced van der Corput set.

Now with help of the well-known Furstenberg corresponding principle (cf. [9, Lemma 2.5]), we can obtain

Corollary 1.1. *Suppose that E is a subset of \mathbb{N} with $\bar{d}_B(E) > 0$ and $\epsilon > 0$. Let*

$$\Lambda_{E,\epsilon}^* = \{p \text{ prime} : \bar{d}_B(E \cap (p - 1 + E)) \geq \bar{d}_B(E)^2 - \epsilon\}.$$

Then for any $m \geq 1$, there exist infinitely many primes $p_0, p_1, \dots, p_m \in \Lambda_{E,\epsilon}^$ with $p_0 < \dots < p_m$ such that $p_m - p_1$ is bounded by a constant C_m .*

Suppose that N is sufficiently large and $W = W_0 \prod_{p \leq w} p$, where w very slowly tends to infinity as $N \rightarrow +\infty$. Let $n \sim N$ mean $N \leq n \leq 2N$. In view of the Maynard sieve method, we need to compute the sum

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n+h) \Omega_n \cdot \mu(A \cap T^{-(n+h-1)} A), \quad (1.1)$$

where $\Omega_n \geq 0$ is some weight and

$$\varpi(n) = \begin{cases} \log n, & \text{if } n \text{ is prime,} \\ 0, & \text{otherwise.} \end{cases}$$

As we shall see later, though seemingly it is not easy to give an asymptotic formula for (1.1), we can obtain a suitable lower bound. Define

$$\mathbf{1}_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{otherwise.} \end{cases}$$

Our strategy is to write

$$\mathbf{1}_A(x) = f_1(x) + f_2(x)$$

under the assumption T is ergodic, where f_1 belongs to $L^2(\mathcal{X}, \mathcal{K}, \mu)$, which is the closed L^2 -subspace generated by all eigenfunctions of T , and f_2 is orthogonal to $L^2(\mathcal{X}, \mathcal{K}, \mu)$.

In Section 3, we shall show that

$$\lim_{N \rightarrow \infty} \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n+h) \Omega_n \int_{\mathcal{X}} f_2 \cdot T^{n+h-1} f_2 d\mu = 0.$$

And in Section 4, we can give a lower bound for

$$\lim_{N \rightarrow \infty} \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n+h) \Omega_n \int_{\mathcal{X}} f_1 \cdot T^{n+h-1} f_1 d\mu,$$

provided that W, b, h satisfy some additional assumptions. In fact, we shall transfer this problem to studying an ergodic transformation on a torus, and use some basic techniques from the Diophantine approximation involving primes. Of course, firstly we need to apply the Maynard sieve method to the exponential sum

$$\lim_{N \rightarrow \infty} \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n+h) \Omega_n \cdot e\left(n \cdot \left(\frac{a}{q} + \theta\right)\right)$$

in Section 2, where as usual let $e(x) = \exp(2\pi\sqrt{-1}x)$ for $x \in \mathbb{R}$. Finally, with help of the ergodic decomposition theorem, we shall conclude the proof of Theorem 1.1 in Section 5.

Throughout the whole paper, let $f(x) \ll g(x)$ mean $f(x) = O(g(x))$ as x tends to $+\infty$, i.e., $|f(x)| \leq C|g(x)|$ for some constant $C > 0$ whenever x is sufficiently large. And \ll_{ϵ} means the implied constant in \ll only depends on ϵ . Let τ and ϕ denote the divisor function and the Euler totient function respectively. In particular, if n is an integer, let $\mu(n)$ denote the value of the Möbius function at n , rather than the measure on \mathcal{X} .

2. MAYNARD'S SIEVE METHOD FOR THE EXPONENTIAL SUMS OVER PRIMES

Let N be a sufficiently large integer and $R = N^{\frac{1}{1000}}$. Suppose that w is a large integer with $w \leq \log \log \log N$. Let

$$W = W_0 \prod_{p \leq w \text{ prime}} p \tag{2.1}$$

where W_0 is a fixed integer to be chosen later.

For distinct integers h_0, h_1, \dots, h_k , we say $\{h_0, h_1, \dots, h_k\}$ is an *admissible* set provided that for any prime p , there exists $1 \leq a \leq p$ satisfying

$$a \not\equiv h_j \pmod{p}$$

for each $0 \leq j \leq k$. We may construct an admissible set whose elements are all the multiple of W_0 . In fact, let

$$h_j = jW_0 \prod_{p \leq k+1} p$$

for $j = 0, 1, \dots, k$. It is easy to see that $\{h_0, h_1, \dots, h_k\}$ is an admissible set.

Suppose that $\{h_0, h_1, \dots, h_k\}$ is an admissible set. We may find $1 \leq b \leq W$ such that

$$b \not\equiv -h_j \pmod{p} \tag{2.2}$$

for any prime $p \leq w$ and each $0 \leq j \leq k$. Further, clearly we may assume that $b \equiv 1 \pmod{W_0}$. Also, assume that w is sufficiently large such that the prime factors of $h_j - h_i$ is not greater than w for each $0 \leq i < j \leq k$. So if $n \equiv b \pmod{W}$, then $W, n + h_0, \dots, n + h_k$ are co-prime.

Throughout this paper, we always make the following assumptions:

- (I) $\{h_0, h_1, \dots, h_k\}$ is an admissible set;
- (II) W, h_0, h_1, \dots, h_k are divisible by W_0 ;
- (III) $W, n + h_0, \dots, n + h_k$ are co-prime for any $n \equiv b \pmod{W}$;
- (IV) $b \equiv 1 \pmod{W_0}$.

Suppose that $F(t_0, t_1, \dots, t_k)$ is a smooth function over \mathbb{R}^{k+1} whose support lies on

$$\{(t_0, t_1, \dots, t_k) : t_0, \dots, t_k \geq 0, t_0 + \dots + t_k \leq 1\}. \quad (2.3)$$

Define

$$\lambda_{d_0, d_1, \dots, d_k}(F) = F\left(\frac{\log d_0}{\log R}, \frac{\log d_1}{\log R}, \dots, \frac{\log d_k}{\log R}\right) \prod_{j=0}^k \mu(d_j)$$

and let

$$\Omega_n(F) = \left(\sum_{\substack{d_i | n + h_i \\ 0 \leq i \leq k}} \lambda_{d_0, d_1, \dots, d_k}(F) \right)^2.$$

The following lemma is the main ingredient of Maynard's sieve method.

Lemma 2.1. *Suppose that $F_1(x_0, \dots, x_k)$ and $F_2(x_0, \dots, x_k)$ are two smooth functions whose supports lie on the area (2.3). Then*

$$\begin{aligned} & \sum_{\substack{d_0, \dots, d_k, e_0, \dots, e_k \\ W, [d_0, e_0], \dots, [d_k, e_k] \text{ coprime}}} \frac{\lambda_{d_0, \dots, d_k}(F_1) \lambda_{e_0, \dots, e_k}(F_2)}{[d_0, e_0] \cdots [d_k, e_k]} \\ &= \frac{1 + o_w(1)}{(\log R)^{k+1}} \cdot \frac{W^{k+1}}{\phi(W)^{k+1}} \int_{\mathbb{R}^{k+1}} \frac{\partial^{k+1} F_1(t_0, \dots, t_k)}{\partial t_0 \cdots \partial t_k} \cdot \frac{\partial^{k+1} F_2(t_0, \dots, t_k)}{\partial t_0 \cdots \partial t_k} dt_0 \cdots dt_k. \end{aligned} \quad (2.4)$$

And (2.4) is also valid if the denominator $[d_0, e_0] \cdots [d_k, e_k]$ in the left side is replaced by $\phi([d_0, e_0] \cdots [d_k, e_k])$.

Proof. See [22, Proposition 5] or [20, Lemma 30]. \square

In this section, we shall establish an analogue of Maynard's sieve method for the exponential sums over primes. Let $o_w(1)$ denote a quantity which tends to 0 as $w \rightarrow +\infty$.

Proposition 2.1. *Let $\epsilon > 0$ be a constant. Suppose that $1 \leq a \leq q \leq N^{\frac{1}{3}-\epsilon} R^{-2}$ with $(a, q) = 1$ and $|\theta| \leq N^{\epsilon-1}$. If $q \mid W$, then for any $0 \leq i \leq k$,*

$$\begin{aligned} & \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) e\left((n + h_i) \left(\frac{a}{q} + \theta\right)\right) \cdot \Omega_n(F) \\ &= e\left(\frac{a(b + h_i)}{q}\right) \cdot \frac{1 + o_w(1)}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \cdot \mathcal{J}_i(F) \sum_{n \sim N} e(n\theta), \end{aligned} \quad (2.5)$$

where

$$\mathcal{J}_i(F) = \int_{\mathbb{R}^k} \left(\frac{\partial^k F(t_0, \dots, t_{i-1}, 0, t_{i+1}, \dots, t_k)}{\partial t_0 \cdots \partial t_{i-1} \partial t_{i+1} \cdots \partial t_k} \right)^2 dt_0 \cdots dt_{i-1} dt_{i+1} \cdots dt_k.$$

Otherwise,

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) e\left((n + h_i) \left(\frac{a}{q} + \theta\right)\right) \cdot \Omega_n \ll_{\epsilon} \frac{1}{w^{1-\epsilon}} \cdot \frac{N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}}. \quad (2.6)$$

Below we shall fix $F(t_0, \dots, t_k)$ as a smooth function whose support lies on (2.3). For convenience, abbreviate $\lambda_{d_0, \dots, d_k}(F)$, $\Omega_n(F)$ and $\mathcal{J}_i(F)$ as $\lambda_{d_0, \dots, d_k}$, Ω_n and \mathcal{J}_i respectively.

The following lemma is an analogue of the Bombieri-Vinogradov theorem for the exponential sums over primes, which was proved by Liu and Zhan [17, Theorem 3].

Lemma 2.2. *For any $A > 0$, there exists $B = B(A) > 0$ such that*

$$\sum_{q \leq K} \max_{(a, q)=1} \max_{|\theta| \leq \delta} \left| \sum_{n \sim x} \varpi(n) e\left(n \left(\frac{a}{q} + \theta\right)\right) - \frac{\mu(q)}{\phi(q)} \sum_{n \sim x} e(n\theta) \right| \ll \frac{x}{\log^A x},$$

where $1 \leq K \leq x^{\frac{1}{3}} \log^{-B} x$ and $\delta = K^{-3} \log^{-B} x$.

Define

$$\mathcal{R}_{q, \delta}(x) = \max_{\substack{1 \leq a \leq q \\ (a, q)=1}} \max_{|\theta| \leq \delta} \left| \sum_{n \sim x} \varpi(n) e\left(n \left(\frac{a}{q} + \theta\right)\right) - \frac{\mu(q)}{\phi(q)} \sum_{n \sim x} e(n\theta) \right|. \quad (2.7)$$

For an assertion P , define

$$\mathbf{1}_P = \begin{cases} 1, & \text{the assertion } P \text{ holds,} \\ 0, & \text{otherwise.} \end{cases}$$

For example, $\mathbf{1}_A(x) = \mathbf{1}_{x \in A}$. Let

$$\mathcal{Z}_q = \{d : (d, q) \text{ and } q/(d, q) \text{ are co-prime}\}.$$

Lemma 2.3. *Suppose that $1 \leq a \leq q$ and $(a, q) = 1$. Let D be a positive integer with $(b, D) = 1$. Then*

$$\begin{aligned} & \sum_{\substack{n \sim x \\ n \equiv b \pmod{D}}} \varpi(n) e\left(n \left(\frac{a}{q} + \theta \right)\right) \\ &= \mu\left(\frac{q}{(D, q)}\right) e\left(\frac{ab \cdot \bar{v}_{D, q}}{(D, q)}\right) \cdot \frac{\mathbf{1}_{D \in \mathcal{Z}_q}}{\phi([D, q])} \sum_{n \sim x} e(n\theta) + O\left(\sum_{t|[D, q]} \frac{(D, t)}{D} \cdot \mathcal{R}_{t, \delta}(x)\right) \end{aligned} \quad (2.8)$$

for any θ with $|\theta| \leq \delta$, where $\bar{v}_{D, q}$ is an integer such that

$$\bar{v}_{D, q} \cdot \frac{q}{(D, q)} \equiv 1 \pmod{(D, q)}$$

if $D \in \mathcal{Z}_q$, and $\bar{v}_{D, q} = 0$ otherwise.

Proof. Clearly

$$\sum_{\substack{n \sim x \\ n \equiv b \pmod{D}}} \varpi(n) e\left(n \left(\frac{a}{q} + \theta \right)\right) = \frac{1}{D} \sum_{r=1}^D e\left(-\frac{br}{D}\right) \sum_{n \sim x} \varpi(n) e\left(n \left(\frac{r}{D} + \frac{a}{q} + \theta \right)\right).$$

Write

$$\frac{r}{D} + \frac{a}{q} = \frac{s_r}{t_r},$$

where t_r is a common multiple of D and q and $(s_r, t_r) = 1$. Then in view of (2.7),

$$\begin{aligned} & \sum_{\substack{n \sim x \\ n \equiv b \pmod{D}}} \varpi(n) e\left(n \left(\frac{a}{q} + \theta \right)\right) \\ &= \frac{1}{D} \sum_{r=1}^D \frac{\mu(t_r)}{\phi(t_r)} \cdot e\left(-\frac{br}{D}\right) \sum_{n \sim x} e(n\theta) + O\left(\sum_{t|[D, q]} \frac{M_t \mathcal{R}_{t, \delta}(x)}{D}\right), \end{aligned} \quad (2.9)$$

where

$$M_t = |\{1 \leq r \leq D : t_r = t\}|.$$

On the other hand, by the Siegel-Walsz theorem, for any $A > 0$, we have

$$\begin{aligned} & \sum_{\substack{n \sim x \\ n \equiv b \pmod{D}}} \varpi(n) e\left(n \cdot \frac{a}{q}\right) = \sum_{\substack{1 \leq c \leq q \\ (c, q) = 1}} e\left(\frac{ac}{q}\right) \sum_{\substack{n \sim x \\ n \equiv b \pmod{D} \\ n \equiv c \pmod{q}}} \varpi(n) \\ &= \frac{x}{\phi([D, q])} \sum_{\substack{1 \leq c \leq q, \\ c \equiv b \pmod{(D, q)}}} e\left(\frac{ac}{q}\right) + O\left(\frac{qx}{(\log x)^A}\right), \end{aligned} \quad (2.10)$$

whenever $x \geq e^{[D,q]}$. Write $u = (D, q)$ and $v = q/u$. We have

$$\begin{aligned} \sum_{\substack{1 \leq c \leq q, (c,q)=1 \\ c \equiv b \pmod{(D,q)}}} e\left(\frac{ac}{q}\right) &= \sum_{d|q} \mu(d) \sum_{\substack{1 \leq c \leq q, d|c \\ c \equiv b \pmod{u}}} e\left(\frac{ac}{q}\right) \\ &= \sum_{d|q} \mu(d) \sum_{\substack{1 \leq t \leq v \\ ut+b \equiv 0 \pmod{d}}} e\left(\frac{a(ut+b)}{q}\right). \end{aligned}$$

Since $(D, b) = 1$ and $u \mid D$, $ut + b \equiv 0 \pmod{d}$ for some integer t if and only if $(u, d) = 1$ and $d \mid v$. It is easy to see that

$$\sum_{\substack{1 \leq t \leq v \\ ut+b \equiv 0 \pmod{d}}} e\left(\frac{a(ut+b)}{q}\right) = 0$$

unless $(u, v) = 1$ and $d = v$. Assume that $(u, v) = 1$. Let \bar{v}_u be an integer with $\bar{v}_u v \equiv 1 \pmod{u}$. When $ut + b \equiv 0 \pmod{v}$,

$$\frac{ut+b}{v} \equiv (ut+b)\bar{v}_u \equiv b\bar{v}_u \pmod{u}.$$

Thus we have

$$\sum_{\substack{1 \leq c \leq q, (c,q)=1 \\ c \equiv b \pmod{(D,q)}}} e\left(\frac{ac}{q}\right) = \begin{cases} \mu(v)e(ab\bar{v}_u/u), & \text{if } (u, v) = 1, \\ 0, & \text{otherwise.} \end{cases}$$

Fix q and D , and let $x \geq e^{[D,q]}$. By Lemma 2.2, we have $\mathcal{R}_{t,\delta}(x) \ll x(\log x)^{-A}$ for any $t \leq [D, q]$ provided that δ is sufficiently small. Thus setting $\theta = 0$ in (2.9) and comparing (2.9) and (2.10), we get

$$\frac{1}{D} \sum_{r=1}^D \frac{\mu(t_r)}{\phi(t_r)} \cdot e\left(-\frac{br}{D}\right) = \frac{\mu(v)e\left(\frac{ab\bar{v}_u}{u}\right) \cdot \mathbf{1}_{(u,v)=1}}{\phi([D, q])}.$$

Suppose that t is a divisor of $[D, q]$ and $t_r = t$. Note that

$$\frac{r}{D} + \frac{a}{q} = \frac{rq + aD}{Dq} = \frac{s_r}{t}.$$

Thus r must satisfy the congruence

$$rq + aD \equiv 0 \pmod{Dq/t}.$$

So

$$M_t \leq \frac{(Dq/t, q)}{Dq/t} \cdot D = (D, t).$$

(2.8) is concluded. \square

We are ready to give the proof of Proposition 2.1. Clearly we only need to consider the case $i = 0$. Let $S_W = \{d \in \mathbb{N} : (d, W) = 1\}$. By Lemma 2.3, we have

$$\begin{aligned}
 & \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) e\left((n + h_0) \left(\frac{a}{q} + \theta\right)\right) \cdot \left(\sum_{d_i | n + h_i} \lambda_{d_0, \dots, d_k}\right)^2 \\
 &= \sum_{\substack{d_1, \dots, d_k, e_1, \dots, e_k \in S_W \\ [d_1, e_1], \dots, [d_k, e_k] \text{ coprime}}} \lambda_{1, d_1, \dots, d_k} \lambda_{1, e_1, \dots, e_k} \sum_{\substack{N + h_0 \leq n \leq 2N + h_0 \\ n \equiv b + h_0 \pmod{W} \\ n \equiv h_0 - h_i \pmod{[d_i, e_i]}}} \varpi(n) e\left(n \left(\frac{a}{q} + \theta\right)\right) \\
 &= \sum_{n \sim N} e(n\theta) \cdot \sum_{\substack{d_1, \dots, d_k, e_1, \dots, e_k \in S_W \\ [d_1, e_1], \dots, [d_k, e_k] \text{ coprime}}} \frac{\lambda_{1, d_1, \dots, d_k} \lambda_{1, e_1, \dots, e_k} \cdot \Delta_{d_1, e_1, \dots, d_k, e_k} \cdot \mathbf{1}_{[W, d_1, e_1, \dots, d_k, e_k] \in \mathcal{Z}_q}}{\phi([W, d_1, e_1, \dots, d_k, e_k, q])} \\
 &+ O\left(\sum_{\substack{d_1, \dots, d_k, e_1, \dots, e_k \in S_W \\ d_1 \cdots d_k, e_1 \cdots e_k \leq R \\ [d_1, e_1], \dots, [d_k, e_k] \text{ coprime}}} \sum_{t | [W, d_1, e_1, \dots, d_k, e_k, q]} \frac{([W, d_1, e_1, \dots, d_k, e_k], t)}{[W, d_1, e_1, \dots, d_k, e_k]} \cdot \mathcal{R}_{t, \delta}(N)\right),
 \end{aligned}$$

where

$$\Delta_{d_1, e_1, \dots, d_k, e_k} = \mu\left(\frac{q}{([W, d_1, e_1, \dots, d_k, e_k], q)}\right) \cdot e\left(\frac{a(b + h_0) \cdot \bar{v}_{[W, d_1, e_1, \dots, d_k, e_k], q}}{([W, d_1, e_1, \dots, d_k, e_k], q)}\right).$$

We firstly consider the remainder term. Assume that $t \leq qWR^2$. We have

$$\sum_{1 \leq D \leq N} \tau(D)^{2k} \cdot \frac{(D, t)}{D} \leq \sum_{s | t} \tau(s)^{2k} \sum_{1 \leq d \leq N/s} \frac{\tau(d)^{2k}}{d} \ll \tau(t)^{2k+1} \cdot (\log N)^{4k},$$

where we use the known result

$$\sum_{1 \leq d \leq x} \tau(d)^k \ll x(\log x)^{2k-1}.$$

It follows from Lemma 2.2 with $K = N^{\frac{1}{3} - \frac{\epsilon}{2}}$ that

$$\begin{aligned}
 & \sum_{\substack{d_1, \dots, d_k, e_1, \dots, e_k \in S_W \\ d_1 \cdots d_k, e_1 \cdots e_k \leq R \\ [d_1, e_1], \dots, [d_k, e_k] \text{ coprime}}} \sum_{t | [W, d_1, e_1, \dots, d_k, e_k, q]} \frac{([W, d_1, e_1, \dots, d_k, e_k], t)}{[W, d_1, e_1, \dots, d_k, e_k]} \cdot \mathcal{R}_{t, \delta}(N) \\
 & \ll \sum_{t \leq N^{\frac{1}{3} - \frac{\epsilon}{2}}} \mathcal{R}_{t, \delta}(N) \sum_{1 \leq D \leq N} \tau(D)^{2k} \cdot \frac{(D, t)}{D} \\
 & \ll (\log N)^{4k} \sum_{t \leq N^{\frac{1}{3} - \frac{\epsilon}{2}}} \tau(t)^{2k+1} \mathcal{R}_{t, \delta}(N) \ll \frac{N}{(\log N)^{k+2}}.
 \end{aligned}$$

Let us turn to the main term. Evidently the main term will vanish if $W \notin \mathcal{Z}_q$. Below we assume that $W \in \mathcal{Z}_q$. Let \mathfrak{X}_q be the set of all those $(\mathfrak{d}_0, \mathfrak{e}_0, \dots, \mathfrak{d}_k, \mathfrak{e}_k)$ satisfying that

- (i) $\mathfrak{d}_0, \mathfrak{e}_0, \dots, \mathfrak{d}_k, \mathfrak{e}_k$ are the divisors of q ;
- (ii) $\mathfrak{d}_0, \mathfrak{e}_0, \dots, \mathfrak{d}_k, \mathfrak{e}_k$ are square-free;
- (iii) $\mathfrak{d}_0, \mathfrak{e}_0, \dots, \mathfrak{d}_k, \mathfrak{e}_k \in \mathcal{Z}_q$;
- (iv) $W, [\mathfrak{d}_0, \mathfrak{e}_0], \dots, [\mathfrak{d}_k, \mathfrak{e}_k]$ are co-prime.

Suppose that $(\mathfrak{d}_0, \mathfrak{e}_0, \dots, \mathfrak{d}_k, \mathfrak{e}_k) \in \mathfrak{X}_q$.

$$\begin{aligned}
& \sum_{\substack{d_1, \dots, d_k, e_1, \dots, e_k \in S_W \\ [d_1, e_1], \dots, [d_k, e_k] \text{ coprime} \\ d_1, \dots, d_k, e_1, \dots, e_k \in \mathcal{Z}_q \\ (d_i, q) = \mathfrak{d}_i, (e_i, q) = \mathfrak{e}_i}} \frac{\lambda_{1, d_1, \dots, d_k} \lambda_{1, e_1, \dots, e_k}}{\phi([d_0, e_0] \cdots [d_k, e_k])} \\
&= \sum_{\substack{(d_1^*, \dots, d_k^*, e_1^*, \dots, e_k^*) \in S_{[W, q]} \\ [d_1^*, e_1^*], \dots, [d_k^*, e_k^*] \text{ coprime}}} \frac{\lambda_{1, \mathfrak{d}_1 d_1^*, \dots, \mathfrak{d}_k d_k^*} \lambda_{1, \mathfrak{e}_1 e_1^*, \dots, \mathfrak{e}_k e_k^*}}{\phi([\mathfrak{d}_1, \mathfrak{e}_1][d_1^*, e_1^*] \cdots [\mathfrak{d}_k, \mathfrak{e}_k][d_k^*, e_k^*])} \\
&= \prod_{j=1}^k \frac{\mu(\mathfrak{d}_j) \mu(\mathfrak{e}_j)}{\phi([\mathfrak{d}_j, \mathfrak{e}_j])} \sum_{\substack{(d_1^*, \dots, d_k^*, e_1^*, \dots, e_k^*) \in S_{[W, q]} \\ W, [d_1^*, e_1^*], \dots, [d_k^*, e_k^*] \text{ coprime}}} \frac{\lambda_{1, d_1^*, \dots, d_k^*}(F_{\mathfrak{d}_1, \dots, \mathfrak{d}_k}) \lambda_{1, e_1^*, \dots, e_k^*}(F_{\mathfrak{e}_1, \dots, \mathfrak{e}_k})}{\phi([d_1^*, e_1^*] \cdots [d_k^*, e_k^*])},
\end{aligned}$$

where

$$F_{\mathfrak{d}_1, \dots, \mathfrak{d}_k}(t_1, \dots, t_k) = F\left(0, t_1 + \frac{\log \mathfrak{d}_1}{\log R}, \dots, t_k + \frac{\log \mathfrak{d}_k}{\log R}\right).$$

According to Lemma 2.1,

$$\begin{aligned}
& \sum_{\substack{(d_1^*, \dots, d_k^*, e_1^*, \dots, e_k^*) \in S_{Wq} \\ W, [d_1^*, e_1^*], \dots, [d_k^*, e_k^*] \text{ are coprime}}} \frac{\lambda_{1, d_1^*, \dots, d_k^*}(F_{\mathfrak{d}_1, \dots, \mathfrak{d}_k}) \lambda_{1, e_1^*, \dots, e_k^*}(F_{\mathfrak{e}_1, \dots, \mathfrak{e}_k})}{\phi([d_1^*, e_1^*] \cdots [d_k^*, e_k^*])} \\
&= \frac{1 + o(1)}{(\log R)^k} \cdot \frac{[W, q]^k}{\phi([W, q])^k} \int_{\mathbb{R}^k} \frac{\partial^k F_{\mathfrak{d}_1, \dots, \mathfrak{d}_k}}{\partial t_1 \cdots \partial t_k} \cdot \frac{\partial^k F_{\mathfrak{e}_1, \dots, \mathfrak{e}_k}}{\partial t_1 \cdots \partial t_k} dt_1 \cdots dt_k.
\end{aligned}$$

Suppose that $q \nmid W$. Letting

$$\Theta_F = \max_{0 \leq t_1, \dots, t_k \leq 1} \left| \frac{\partial^k F(0, t_1, \dots, t_k)}{\partial t_1 \cdots \partial t_k} \right|,$$

we get

$$\left| \sum_{\substack{d_1, \dots, d_k, e_1, \dots, e_k \in S_W \\ [d_1, e_1], \dots, [d_k, e_k] \text{ coprime} \\ d_1, \dots, d_k, e_1, \dots, e_k \in \mathcal{Z}_q \\ (d_i, q) = \mathfrak{d}_i, (e_i, q) = \mathfrak{e}_i}} \frac{\lambda_{1, d_1, \dots, d_k} \lambda_{1, e_1, \dots, e_k}}{\phi([d_0, e_0] \cdots [d_k, e_k])} \right| \leq \prod_{j=1}^k \frac{1}{\phi([\mathfrak{d}_j, \mathfrak{e}_j])} \cdot \frac{1 + o_w(1)}{(\log R)^k} \cdot \frac{[W, q]^k}{\phi([W, q])^k} \cdot \Theta_F^2.$$

Thus

$$\begin{aligned} & \sum_{\substack{d_1, \dots, d_k, e_1, \dots, e_k \in S_W \\ d_1, e_1, \dots, d_k, e_k \in \mathcal{Z}_q \\ [d_1, e_1], \dots, [d_k, e_k] \text{ coprime}}} \Delta_{d_1, e_1, \dots, d_k, e_k} \cdot \frac{\lambda_{1, d_1, \dots, d_k} \lambda_{1, e_1, \dots, e_k}}{\phi([W, d_1, e_1, \dots, d_k, e_k, q])} \\ &= \sum_{\mathfrak{d}_1, \mathfrak{e}_1, \dots, \mathfrak{d}_k, \mathfrak{e}_k \in \mathfrak{X}_q} \frac{\Delta_{d_1, e_1, \dots, d_k, e_k}}{\phi\left(\frac{[W, q]}{[\mathfrak{d}_1, \mathfrak{e}_1] \cdots [\mathfrak{d}_k, \mathfrak{e}_k]}\right)} \sum_{\substack{d_1, \dots, d_k, e_1, \dots, e_k \in S_W \\ [d_1, e_1], \dots, [d_k, e_k] \text{ are coprime} \\ d_1, \dots, d_k, e_1, \dots, e_k \in \mathcal{Z}_q \\ (d_i, q) = \mathfrak{d}_i, (e_i, q) = \mathfrak{e}_i}} \frac{\lambda_{1, d_1, \dots, d_k} \lambda_{1, e_1, \dots, e_k}}{\phi([d_1, e_1] \cdots [d_k, e_k])} \\ &\ll \sum_{\mathfrak{d}_1, \mathfrak{e}_1, \dots, \mathfrak{d}_k, \mathfrak{e}_k \in \mathfrak{X}_q} \frac{1}{\phi\left(\frac{[W, q]}{[\mathfrak{d}_1, \mathfrak{e}_1] \cdots [\mathfrak{d}_k, \mathfrak{e}_k]}\right)} \cdot \prod_{j=1}^k \frac{1}{\phi([\mathfrak{d}_j, \mathfrak{e}_j])} \cdot \frac{1}{(\log R)^k} \cdot \frac{[W, q]^k}{\phi([W, q])^k} \cdot \Theta_F^2 \\ &= \frac{\Theta_F^2}{(\log R)^k} \cdot \frac{[W, q]^k}{\phi([W, q])^{k+1}} \sum_{\mathfrak{d}_1, \mathfrak{e}_1, \dots, \mathfrak{d}_k, \mathfrak{e}_k \in \mathfrak{X}_q} 1. \end{aligned}$$

Let $q_* = q/(W, q)$. Since $q \nmid W$ and $W \in \mathcal{Z}_q$, q_* must have at least one prime factor greater than w , i.e., $q_* > w$. And $\mathfrak{d}_1, \mathfrak{e}_1, \dots, \mathfrak{d}_k, \mathfrak{e}_k \in \mathfrak{X}_q$ implies that $\mathfrak{d}_1, \mathfrak{e}_1, \dots, \mathfrak{d}_k, \mathfrak{e}_k$ all divide q_* . Hence

$$\frac{[W, q]^k}{\phi([W, q])^{k+1}} \sum_{\mathfrak{d}_1, \mathfrak{e}_1, \dots, \mathfrak{d}_k, \mathfrak{e}_k \in \mathfrak{X}_q} 1 \leq \frac{W^k}{\phi(W)^{k+1}} \cdot \frac{\tau(q_*)^{2k} q_*^k}{\phi(q_*)^{k+1}} \ll_{\epsilon} \frac{W^k}{\phi(W)^{k+1}} \cdot \frac{1}{w^{1-\epsilon}}.$$

Finally, if $q \mid W$, then evidently $\mathfrak{X}_q = \{(1, 1, \dots, 1)\}$. So

$$\begin{aligned} & \sum_{\substack{d_1, \dots, d_k, e_1, \dots, e_k \in S_W \\ d_1, e_1, \dots, d_k, e_k \in \mathcal{Z}_q \\ [d_1, e_1], \dots, [d_k, e_k] \text{ coprime}}} \Delta_{d_1, e_1, \dots, d_k, e_k} \cdot \frac{\lambda_{1, d_1, \dots, d_k} \lambda_{1, e_1, \dots, e_k}}{\phi([W, d_1, e_1, \dots, d_k, e_k, q])} \\ &= e\left(\frac{a(b + h_0)}{q}\right) \cdot \frac{1}{\phi(W)} \sum_{\substack{d_1, \dots, d_k, e_1, \dots, e_k \in S_W \\ [d_1, e_1], \dots, [d_k, e_k] \text{ are coprime}}} \frac{\lambda_{1, d_1, \dots, d_k} \lambda_{1, e_1, \dots, e_k}}{\phi([d_1, e_1] \cdots [d_k, e_k])} \\ &= e\left(\frac{a(b + h_0)}{q}\right) \cdot \frac{1 + o_w(1)}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \int_{\mathbb{R}^k} \left(\frac{\partial^k F(0, t_1, \dots, t_k)}{\partial t_1 \cdots \partial t_k} \right)^2 dt_1 \cdots dt_k. \end{aligned}$$

□

Remark 2.1. If we set $a = q = 1$ and $\theta = 0$ in (2.5), we can get

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) \cdot \Omega_n = (1 + o_w(1)) \cdot \frac{\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}}, \quad (2.11)$$

which is one of two key formulas used in the proof of the Maynard-Tao theorem. And the other one is

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \Omega_n = (1 + o_w(1)) \mathcal{J}_* \cdot \frac{N}{(\log R)^{k+1}} \cdot \frac{W^k}{\phi(W)^{k+1}}, \quad (2.12)$$

where

$$\mathcal{J}_* = \int_{\mathbb{R}^k} \left(\frac{\partial^{k+1} F(t_0, \dots, t_k)}{\partial t_0 \cdots \partial t_k} \right)^2 dt_0 \cdots dt_k.$$

3. REDUCING TO THE KRONECKER SYSTEM

Suppose that $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, T)$ is a measure-preserving system and T is invertible. For any $f(x), g(x) \in L^2(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu)$, let

$$\langle f, g \rangle := \int_{\mathcal{X}} f \cdot \bar{g} d\mu$$

denote the inner product over $L^2(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu)$. In particular, let $\|f\|_2 := \sqrt{\langle f, f \rangle}$ denote the L^2 -norm of f .

Let $\mathcal{K} \subseteq \mathcal{B}_{\mathcal{X}}$ be the smallest sub- σ -algebra with respect to which all eigenfunctions of T are measurable. For any $f(x) \in L^2_{\nu}$, let $\mathbb{E}(f|\mathcal{K})$ denote the conditional expectation of f with respect to \mathcal{K} (cf. [7, Theorem 5.1]). It is known that

$$\int_A f d\mu = \int_A \mathbb{E}(f|\mathcal{K}) d\mu$$

for any $A \in \mathcal{K}$. And if f is non-negative, then $\mathbb{E}(f|\mathcal{K})$ is also non-negative almost everywhere. In fact, $\mathbb{E}(f|\mathcal{K})$ is the orthogonal projection of f over the close Hilbert subspace $L^2(\mathcal{X}, \mathcal{K}, \mu)$. Write $f_1 = \mathbb{E}(f|\mathcal{K})$ and $f_2 = f - \mathbb{E}(f|\mathcal{K})$. Then f_2 is orthogonal to $L^2(\mathcal{X}, \mathcal{K}, \mu)$. Since $T^{-1}(\mathcal{K}) \subseteq \mathcal{K}$, $T^n f_1 \in L^2(\mathcal{X}, \mathcal{K}, \mu)$ for any $n \in \mathbb{N}$. Thus we have

$$\langle f, T^n f \rangle = \langle f_1, T^n f_1 \rangle + \langle f_2, T^n f_2 \rangle.$$

In this section, we shall focus on those $f(x)$ with $\mathbb{E}(f|\mathcal{K}) = 0$.

Proposition 3.1. *Suppose that $\mathbb{E}(f|\mathcal{K}) = 0$. Then for each $0 \leq i \leq k$,*

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) \Omega_n \cdot \langle f, T^{n+h_i-1} f \rangle = o_w \left(\frac{N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \right), \quad (3.1)$$

whenever N is sufficiently large with respect to w .

Clearly we only need to consider the case $i = 0$. Let $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ denote the 1-dimensional torus. For convenience, let $|\cdot|_{\mathbb{T}}$ denote the norm over \mathbb{T} , i.e., for any $x \in \mathbb{R}$

$$|x|_{\mathbb{T}} := \min\{|x - n| : n \in \mathbb{Z}\}.$$

Let $P = N^{\frac{1}{3} - \frac{1}{99}}$ and $Q = N^{1 - \frac{1}{49}}$. Define

$$\mathfrak{M}_N^{(a,q)} = \{\alpha \in \mathbb{T} : |\alpha - a/q|_{\mathbb{T}} \leq q^{-1}Q^{-1}\}.$$

By a well-known result of Dirichlet, we have

$$\mathbb{T} = \bigcup_{\substack{1 \leq a \leq q \leq Q \\ (a,q)=1}} \mathfrak{M}_N^{(a,q)}.$$

Let

$$\mathfrak{M}_N = \bigcup_{\substack{1 \leq a \leq q \leq P \\ (a,q)=1}} \mathfrak{M}_N^{(a,q)}$$

and let $\mathfrak{m}_N = \mathbb{T} \setminus \mathfrak{M}_N$.

By the well-known Herglotz theorem (cf. [13, Chapter 1.8]), there exists a non-negative measure ν on \mathbb{T} such that

$$\langle f, T^n f \rangle = \int_0^1 e(n\alpha) d\nu(\alpha).$$

First, we consider the integrals on the minor arc \mathfrak{m}_N . The following lemma is due to Balog and Perelli [1].

Lemma 3.1. *Suppose that $(a, q) = (b, D) = 1$. Then letting $u_D = (D, q)$, we have*

$$\sum_{\substack{n \leq x \\ n \equiv b \pmod{D}}} \varpi(n) e\left(n \cdot \frac{a}{q}\right) \ll (\log x)^3 \left(\frac{u_D x}{D q^{\frac{1}{2}}} + \frac{q^{\frac{1}{2}} x^{\frac{1}{2}}}{u_D^{\frac{1}{2}}} + \frac{x^{\frac{4}{5}}}{D^{\frac{2}{5}}} \right). \quad (3.2)$$

Applying Lemma 3.1 and the partial summation, we can get

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{D}}} \varpi(n) e\left(n \left(\frac{a}{q} + \theta \right)\right) \ll (1 + \theta N) (\log N)^3 \cdot \left(\frac{u_D N}{D q^{\frac{1}{2}}} + \frac{q^{\frac{1}{2}} N^{\frac{1}{2}}}{u_D^{\frac{1}{2}}} + \frac{N^{\frac{4}{5}}}{D^{\frac{2}{5}}} \right),$$

where $u_D = (D, q)$. It follows that

$$\begin{aligned}
& \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) e\left((n + h_0) \left(\frac{a}{q} + \theta\right)\right) \cdot \Omega_n \\
& \ll \sum_{\substack{d_1, \dots, d_k, e_1, \dots, e_k \in S_W \\ d_1 \dots d_k, e_1 \dots e_k \leq R \\ [d_1, e_1], \dots, [d_k, e_k] \text{ coprime}}} \left| \sum_{\substack{n \sim N \\ n \equiv b + h_0 \pmod{W} \\ n \equiv h_0 - h_i \pmod{[d_i, e_i]}}} \varpi(n) e\left(n \left(\frac{a}{q} + \theta\right)\right) \right| \\
& \ll (1 + \theta N)(\log N)^3 \sum_{D \leq WR^2} \tau(D)^{2k} \cdot \left(\frac{u_D N}{D q^{\frac{1}{2}}} + \frac{q^{\frac{1}{2}} N^{\frac{1}{2}}}{u_D^{\frac{1}{2}}} + \frac{N^{\frac{4}{5}}}{D^{\frac{2}{5}}} \right).
\end{aligned}$$

Note that

$$\begin{aligned}
\sum_{D \leq WR^2} \tau(D)^{2k} \cdot \frac{u_D}{D} & \leq \sum_{u|q} \sum_{\substack{D \leq WR^2 \\ u|D}} \frac{\tau(D)^{2k}}{D/u} \leq \sum_{u|q} \tau(u)^{2k} \sum_{v \leq WR^2/u} \frac{\tau(v)^{2k}}{v} \\
& \ll (\log N)^{4k} \sum_{u|q} \tau(u)^{2k} \ll (\log N)^{4k} \cdot \tau(q)^{2k+1}.
\end{aligned}$$

Hence for any $\epsilon > 0$,

$$\begin{aligned}
& \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) e\left((n + h_0) \left(\frac{a}{q} + \theta\right)\right) \cdot \Omega_n \\
& \ll_{\epsilon} \frac{N^{1+\epsilon}}{q^{\frac{1}{2}-\epsilon}} + \frac{\theta N^{2+\epsilon}}{q^{\frac{1}{2}-\epsilon}} + q^{\frac{1}{2}} N^{\frac{1}{2}+\epsilon} R^2 + \theta q^{\frac{1}{2}} N^{\frac{3}{2}+\epsilon} R^2 + N^{\frac{4}{5}+\epsilon} R^{\frac{6}{5}} + \theta N^{\frac{9}{5}+\epsilon} R^{\frac{6}{5}}. \quad (3.3)
\end{aligned}$$

Suppose that $\alpha \in \mathfrak{m}_N$. Then $\alpha = a/q + \theta$ where $P < q \leq Q$, $(a, q) = 1$ and $|\theta| \leq q^{-1}Q^{-1}$. By (3.3),

$$\begin{aligned}
& \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) e((n + h_0)\alpha) \cdot \Omega_n \\
& \ll_{\epsilon} \frac{N^{1+\frac{1}{99}}}{q^{\frac{1}{2}-\frac{1}{99}}} + \frac{N^{2+\frac{1}{99}}}{q^{\frac{3}{2}-\frac{1}{99}}Q} + q^{\frac{1}{2}} N^{\frac{1}{2}+\frac{1}{99}} R^2 + \frac{N^{\frac{3}{2}+\frac{1}{99}} R^2}{q^{\frac{1}{2}}Q} + N^{\frac{4}{5}+\frac{1}{99}} R^{\frac{6}{5}} + \frac{N^{\frac{9}{5}+\frac{1}{99}} R^{\frac{6}{5}}}{qQ} \\
& \leq \frac{N^{1+\frac{1}{99}}}{P^{\frac{1}{2}-\frac{1}{99}}} + \frac{N^{2+\frac{1}{99}}}{P^{\frac{3}{2}-\frac{1}{99}}Q} + Q^{\frac{1}{2}} N^{\frac{1}{2}+\frac{1}{99}} R^2 + \frac{N^{\frac{3}{2}+\frac{1}{99}} R^2}{P^{\frac{1}{2}}Q} + N^{\frac{4}{5}+\frac{1}{99}} R^{\frac{6}{5}} + \frac{N^{\frac{9}{5}+\frac{1}{99}} R^{\frac{6}{5}}}{PQ} \\
& \ll N^{1-\frac{1}{999}}. \quad (3.4)
\end{aligned}$$

Hence

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \int_{\mathfrak{m}_N} e((n + h_0 - 1)\alpha) d\nu(\alpha) \ll N^{1-\frac{1}{999}}.$$

Next, let us turn to the integrals on \mathfrak{M}_N . Since $\mathbb{E}(f|\mathcal{K}) = 0$, we must have $f(x)$ is orthogonal to any eigenfunction of T . It follows that v is non-atomic, i.e., for any $\theta \in \mathbb{T}$,

$$\lim_{\epsilon \rightarrow 0} \int_{\theta-\epsilon}^{\theta+\epsilon} 1dv = 0.$$

For any $k \geq 1$, Let X_k be the least positive integer such that if $N \geq X_k$, then

$$\sum_{\substack{1 \leq a \leq q \leq k \\ (a,q)=1}} \int_{\mathfrak{M}_N^{(a,q)}} 1dv \leq \frac{1}{k}.$$

Define $\Psi(x) = k$ if $X_k \leq x < X_{k+1}$. Then $\Psi(x)$ tends to infinity as $x \rightarrow +\infty$, and

$$\lim_{N \rightarrow +\infty} \sum_{\substack{1 \leq a \leq q \leq \Psi(N) \\ (a,q)=1}} \int_{\mathfrak{M}_N^{(a,q)}} 1dv \leq \lim_{N \rightarrow +\infty} \frac{1}{\Psi(N)} = 0.$$

Assume that N is sufficiently large such that

$$w \leq \frac{1}{3} \log \Psi(N).$$

Then

$$W = W_0 \prod_{p \leq w} p \leq \Psi(N).$$

Let

$$\mathfrak{M}_N^* = \bigcup_{\substack{1 \leq a \leq q \leq \Psi(N) \\ (a,q)=1}} \mathfrak{M}_N^{(a,q)}$$

Then by (2.5),

$$\begin{aligned} & \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n+h_0) \Omega_n \int_{\mathfrak{M}_N^*} e((n+h_0-1)\alpha) dv(\alpha) \\ &= \sum_{\substack{1 \leq a \leq q \leq \Psi(N) \\ (a,q)=1}} \int_{\mathfrak{M}_N^{(a,q)}} \left(\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n+h_0) \Omega_n \cdot e((n+h_0-1)\alpha) \right) dv(\alpha) \\ &\ll \frac{N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \sum_{\substack{1 \leq a \leq q \leq \Psi(N) \\ (a,q)=1}} \int_{\mathfrak{M}_N^{(a,q)}} 1dv = o\left(\frac{N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}}\right). \end{aligned}$$

On the other hand, if $\mathfrak{M}_N^{(a,q)} \subseteq \mathfrak{M}_N \setminus \mathfrak{M}_N^*$, then we must have $q > \Psi(N) \geq W$, i.e., $q \nmid W$. It follows from (2.6) that

$$\begin{aligned}
& \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \int_{\mathfrak{M}_N \setminus \mathfrak{M}_N^*} e((n + h_0 - 1)\alpha) dv(\alpha) \\
&= \sum_{\substack{\Psi(N) < q \leq P \\ 1 \leq a \leq q, (a,q)=1}} \int_{\mathfrak{M}_N^{(a,q)}} \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \cdot e((n + h_0 - 1)\alpha) dv(\alpha) \\
&\ll \frac{1}{\sqrt{w}} \cdot \frac{N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \cdot \sum_{\substack{\Psi(N) < q \leq P \\ 1 \leq a \leq q, (a,q)=1}} \int_{\mathfrak{M}_N^{(a,q)}} 1 dv \\
&= o_w \left(\frac{N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \right).
\end{aligned}$$

Thus

$$\begin{aligned}
& \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \cdot \langle f, T^{n+h_0-1} f \rangle \\
&= \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \int_0^1 e((n + h_0 - 1)\alpha) dv(\alpha) \\
&= o_w \left(\frac{N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \right).
\end{aligned}$$

□

4. ERGODIC ROTATION ON A TORUS

In this section, we assume that $f(x) \in L^2(\mathcal{X}, \mathcal{K}, \mu)$. Let

$$\|f\|_1 := \int_{\mathcal{X}} |f| d\mu$$

denote the L^1 -norm of f .

Lemma 4.1. *There exist a measure-preserving system $(\mathcal{G}, \mathcal{B}_{\mathcal{G}}, \nu, S_{\alpha})$ and a map ψ from \mathcal{X} to \mathcal{G} such that*

- (1) \mathcal{G} is a compact abelian group and ν is the Haar measure on \mathcal{G} ;
- (2) $S_{\alpha} : x \mapsto x + \alpha$ is an ergodic rotation on \mathcal{G} , where $\alpha \in \mathcal{G}$;
- (3) $\psi \circ T = S_{\alpha} \circ \psi$ and $\nu = \mu \circ \psi^{-1}$;
- (4) $\mathcal{K} = \psi^{-1}(\mathcal{B}_{\mathcal{G}}) \pmod{\mu}$;

Proof. See [7, Theorem 6.10]. \square

Here $(\mathcal{G}, \mathcal{B}_{\mathcal{G}}, \nu, S_{\alpha})$ is also called the *Kronecker factor* of $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, T)$. Since \mathcal{G} is a compact abelian group, \mathcal{G} is isomorphic to the direct sum of a finite abelian group and a torus. So we may assume that $\mathcal{G} = G \oplus \mathbb{T}^d$ and $\nu = \nu_G \times \mathbf{m}_{\mathbb{T}^d}$, where ν_G is the discrete probability measure on G and $\mathbf{m}_{\mathbb{T}^d}$ is the Lebesgue measure on \mathbb{T}^d . Below we may assume that the integer W_0 in (2.1) is divisible by the cardinality of G .

Proposition 4.1. *Suppose that $f(x) \in L^2(\mathcal{X}, \mathcal{K}, \mu)$ is a non-negative function with $0 < \|f\|_2 \leq 1$ and $0 < \epsilon < \|f\|_1^2$. Then for each $0 \leq i \leq k$*

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) \Omega_n \cdot \langle f, T^{n+h_i-1} f \rangle \geq (\|f\|_1^2 - \epsilon) \cdot \frac{\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}}, \quad (4.1)$$

whenever N is sufficiently large with respect to w .

It suffices to prove Proposition 4.1 when $i = 0$. Let $\epsilon_1 = \epsilon/12$. Choose disjoint $A_1, A_2, \dots, A_s \in \mathcal{K}$ and $a_1, \dots, a_s > 0$ such that

$$\|f - f_*\|_2 \leq \epsilon_1 \|f\|_1,$$

where

$$f_*(x) := \sum_{i=1}^s a_i \cdot \mathbf{1}_{A_i}(x).$$

Suppose that $B_1, B_2, \dots, B_s \in \mathcal{B}_{\mathcal{G}}$ and $\psi^{-1}(B_i) = A_i$ for each $1 \leq i \leq s$. Let

$$g_*(x) := \sum_{i=1}^s \alpha_i \cdot \mathbf{1}_{B_i}(x).$$

Then $f = g_* \circ \psi$. Since $\nu = \mu \circ \psi^{-1}$, we have $\|f_*\|_2 = \|g_*\|_2$, as well as $\|f_*\|_1 = \|g_*\|_1$. We shall approximate those B_1, \dots, B_s by using some small d -cubes on \mathbb{T}^d . Define the d -cube

$$\mathfrak{C}_{\mathbf{a}, \eta} = \{(t_1, \dots, t_d) : a_i \eta \leq t_i < (a_i + 1)\eta \text{ for } i = 1, \dots, d\}.$$

where $\mathbf{a} = (a_1, \dots, a_d) \in \mathbb{Z}^d$ and $0 \leq a_1, \dots, a_d \leq \eta^{-1} - 1$. Trivially $\mathbf{m}_{\mathbb{T}^d}(\mathfrak{C}_{\mathbf{a}, \eta}) = \eta^d$.

Choose a sufficiently small constant $\eta_0 > 0$, $\beta_1, \dots, \beta_t > 0$ and $E_1, \dots, E_t \in \mathcal{B}_{\mathcal{G}}$ such that

(a) the L^2 -norm

$$\|g_* - g\|_2 \leq \epsilon_1 \|g_*\|_1,$$

where

$$g(x) := \sum_{i=1}^t \beta_i \cdot \mathbf{1}_{E_i}(x);$$

(b) each E_j is of the form $E_j = (\gamma_j, \mathfrak{C}_{\mathbf{a}_j, \eta_0})$, where $\gamma_j \in H$.

Thus we have

$$\begin{aligned}\langle f, T^n f \rangle &\geq \langle f_*, T^n f_* \rangle - 2\|f - f_*\|_2 \cdot \|f\|_2 - 3\|f - f_*\|_2^2 \\ &= \langle f_*, T^n f_* \rangle - 3\epsilon_1 = \langle g_*, S_\alpha^n g_* \rangle - 3\epsilon_1 \geq \langle g, S_\alpha^n g \rangle - 6\epsilon_1.\end{aligned}$$

for any $n \in \mathbb{N}$. That is,

$$\begin{aligned}&\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \cdot \langle f, T^{n+h_0-1} f \rangle \\ &\geq \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \cdot \langle g, S_\alpha^{n+h_0-1} g \rangle - 6\epsilon_1 \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n.\end{aligned}$$

Since S_α is ergodic rotation on $\mathcal{G} = G \oplus \mathbb{T}^d$, we must have

$$\alpha = (\gamma_0, \kappa_1, \dots, \kappa_d),$$

where $\gamma_0 \in G$ and $\kappa_1, \dots, \kappa_d \in \mathbb{T}$ are \mathbb{Q} -linearly independent. Let $L_0 \in \mathbb{N}$ be the least integer such that

$$\left(1 - \frac{1}{L_0}\right)^d > \left(1 - \frac{\epsilon_1}{\|g\|_1^2}\right)^{\frac{1}{3}}.$$

Define the rotation $S_{\kappa_1, \dots, \kappa_d}$ over \mathbb{T}^d by $S_{\kappa_1, \dots, \kappa_d}(x_1, \dots, x_d) = (x_1 + \kappa_1, \dots, x_d + \kappa_d)$.

Lemma 4.2. *Suppose that $\mathbf{a} = (a_1, \dots, a_d)$ and $\mathbf{b} = (b_1, \dots, b_d)$, where $a_1, \dots, a_d, b_1, \dots, b_d$ are integers with $0 \leq a_1, \dots, a_d, b_1, \dots, b_d \leq \eta_0^{-1} - 1$. Then*

$$\begin{aligned}&\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \cdot \mathbf{m}_{\mathbb{T}^d}(\mathfrak{C}_{\mathbf{a}, \eta_0} \cap S_{\kappa_1, \dots, \kappa_d}^{-(n+h_0-1)} \mathfrak{C}_{\mathbf{b}, \eta_0}) \\ &\geq \left(1 - \frac{1}{L_0}\right)^{3d} \cdot \eta_0^{2d} \cdot \frac{\mathcal{J}_0 N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}},\end{aligned}\tag{4.2}$$

whenever N is sufficiently large with respect to w .

Proof. Let

$$\delta_0 = \frac{\eta_0}{L_0}.$$

For $2 - L_0 \leq u_1, u_2, \dots, u_d \leq L_0 - 2$, for each $1 \leq i \leq d$, it is easy to check that

$$m_{\mathbb{T}^d}(\mathfrak{C}_{\mathbf{a}, \eta_0} \cap S_{\kappa_1, \dots, \kappa_d}^{-n} \mathfrak{C}_{\mathbf{b}, \eta_0}) \geq \delta_0^d \prod_{i=1}^d (L_0 - |u_i| - \mathbf{1}_{u_i \geq 0}),$$

provided that

$$a_i \eta_0 + u_i \delta_0 \leq |b_i \eta_0 - n \kappa_i|_{\mathbb{T}} \leq a_i \eta_0 + (u_i + 1) \delta_0.\tag{4.3}$$

Let

$$\delta_1 = \frac{\delta_0}{L_0}.$$

Let $\psi(x)$ be a smooth function on \mathbb{R} with the period 1 such that

- (1) $0 \leq \psi(x) \leq 1$ for any $x \in \mathbb{R}$;
- (2) $\psi(x) = 1$ if $\delta_1 \leq x \leq \delta_0 - \delta_1$, and $\psi(x) = 0$ if $\delta_0 \leq x < 1$;
- (3) $\psi(x)$ has the Fourier expansion

$$\psi(x) = (\delta_0 - \delta_1) + \sum_{|j| \geq 1} \alpha_j e(jx),$$

where

$$\alpha_j \ll C_0 \cdot \min\{|j|^{-1}, \delta_0 - \delta_1, \delta_1^{-1}|j|^{-2}\}$$

for some constant $C_0 > 0$.

It is well-known that such $\psi(x)$ always exists (cf. [23, Lemma 12 of Chapter I]). Let $K_0 \in \mathbb{N}$ be the least integer such that

$$\frac{(\log K_0 + \delta_1^{-1} K_0^{-1} + 1)^{d-1}}{K_0} < \frac{(\delta_0 - \delta_1)^d \delta_1}{2^d C_0^d d} \cdot \left(1 - \left(1 - \frac{1}{L_0}\right)^d\right).$$

Fix u_1, u_2, \dots, u_d with $|u_1|, \dots, |u_d| \leq L_0 - 2$. Let $\Delta_i = (b_i - a_i)\eta_0 - (u_i + 1)\delta_0$. Clearly $\psi((n + h_0 - 1)\kappa_i - \Delta_i) > 0$ implies (4.3) holds. We have

$$\begin{aligned} & \prod_{i=1}^d \psi((n + h_0 - 1)\kappa_i - \Delta_i) \\ &= \prod_{i=1}^d \left((\delta_0 - \delta_1) + \sum_{|j| \geq 1} \alpha_j e(j((n + h_0 - 1)\kappa_i - \Delta_i)) \right) \\ &= (\delta_0 - \delta_1)^d + \sum_{\emptyset \neq I \subseteq \{1, \dots, d\}} \prod_{i \in I} \left(\sum_{1 \leq |j| \leq K_0} \alpha_j e(j((n + h_0 - 1)\kappa_i - \Delta_i)) \right) + \mathcal{R}, \end{aligned}$$

where

$$\begin{aligned} |\mathcal{R}| &\leq d \cdot \left(\sum_{|j| > K_0} \frac{C_0}{\delta_1 j^2} \right) \cdot \left((\delta_0 - \delta_1) + \sum_{1 \leq |j| \leq K_0} \frac{C_0}{j} + \sum_{|j| > K_0} \frac{C_0}{\delta_1 j^2} \right)^{d-1} \\ &\leq \frac{2^d C_0^d d (\log K_0 + \delta_1^{-1} K_0^{-1} + 1)^{d-1}}{\delta_1 K_0} < (\delta_0 - \delta_1)^d \cdot \left(1 - \left(1 - \frac{1}{L_0}\right)^d\right). \end{aligned}$$

Hence

$$\begin{aligned}
& \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \prod_{i=1}^d \psi((n + h_0 - 1)\kappa_i - \Delta_i) \\
& \geq (\delta_0 - \delta_1)^d \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n - |\mathcal{R}| \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \\
& \quad - \sum_{\emptyset \neq I \subseteq \{1, \dots, d\}} \sum_{\substack{i \in I \\ 1 \leq |j_i| \leq K_0}} \prod_{i \in I} |\alpha_{j_i}| \cdot \left| \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \cdot e\left(n \sum_{i \in I} j_i \kappa_i\right) \right|.
\end{aligned}$$

Assume that $I = \{i_1, \dots, i_l\}$ is a non-empty subset of $\{1, \dots, d\}$ and j_{i_1}, \dots, j_{i_l} are integers with $1 \leq |j_{i_1}|, \dots, |j_{i_l}| \leq K_0$. Let

$$\vartheta_{I, j_{i_1}, \dots, j_{i_l}} = \sum_{i \in I} j_i \kappa_i.$$

Since $\kappa_1, \dots, \kappa_d$ are \mathbb{Q} -linearly independent, $\vartheta_{I, j_{i_1}, \dots, j_{i_l}}$ is also irrational. For any $x \geq 1$, there exists $1 \leq a \leq q \leq x$ with $(a, q) = 1$ such that

$$\left| \vartheta_{I, j_{i_1}, \dots, j_{i_l}} - \frac{a}{q} \right|_{\mathbb{T}} \leq \frac{1}{qx}.$$

Let $\varrho_{I, j_{i_1}, \dots, j_{i_l}}(x)$ be the least one of such q 's. Clearly $\varrho_{I, j_{i_1}, \dots, j_{i_l}}(x)$ is an increasing function tending to ∞ as $x \rightarrow +\infty$. Define

$$\varrho(x) = \min_{\substack{\emptyset \neq I \subseteq \{1, \dots, d\} \\ I = \{i_1, \dots, i_l\} \\ 1 \leq |j_{i_1}|, \dots, |j_{i_l}| \leq K_0}} \varrho_{I, j_{i_1}, \dots, j_{i_l}}(x)$$

Assume that N is sufficiently large such that

$$w < \frac{1}{3} \log \varrho(N),$$

i.e., $W < \varrho(N)$. Let $P = N^{\frac{1}{3} - \frac{1}{99}}$ and $Q = N^{1 - \frac{1}{49}}$. For any $\vartheta_{I, j_{i_1}, \dots, j_{i_l}}$, there exists $\varrho(N) \leq q \leq Q$ and $1 \leq a \leq q$ with $(a, q) = 1$ such that

$$\left| \vartheta_{I, j_{i_1}, \dots, j_{i_l}} - \frac{a}{q} \right|_{\mathbb{T}} \leq \frac{1}{qQ}.$$

If $q \geq P$, according to (3.4),

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) e(\vartheta_{I, j_{i_1}, \dots, j_{i_l}} n) \cdot \Omega_n \ll N^{1 - \frac{1}{999}}.$$

Suppose that $\varrho(N) \leq q < P$. Since $q \nmid W$, using (2.6), we also have

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \cdot e(n \vartheta_{I, j_{i_1}, \dots, j_{i_l}}) = o_w \left(\frac{N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \right).$$

Therefore, in view of (2.11), we get

$$\begin{aligned} & \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \prod_{i=1}^d \psi((n + h_0 - 1) \kappa_i - \Delta_i) \\ & \geq ((\delta_0 - \delta_1)^d - |\mathcal{R}|) \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n + o_w \left(\frac{N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \right) \\ & \geq \left(1 - \frac{1}{L_0} \right)^d \cdot (\delta_0 - \delta_1)^d \cdot \frac{\mathcal{J}_0 N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \\ & = \left(1 - \frac{1}{L_0} \right)^{2d} \cdot \delta_0^d \cdot \frac{\mathcal{J}_0 N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}}, \end{aligned}$$

provided that w is sufficiently large.

On the other hand,

$$\sum_{|u_1|, \dots, |u_d| \leq L_0 - 2} \prod_{i=1}^d (L_0 - |u_i| - \mathbf{1}_{u_i \geq 0}) = \left(\sum_{u=2-L_0}^{L_0-2} (L_0 - |u| - \mathbf{1}_{u \geq 0}) \right)^d = L_0^d (L_0 - 1)^d.$$

Thus

$$\begin{aligned} & \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \cdot \mathbf{m}_{\mathbb{T}^d}(\mathfrak{C}_{\mathbf{a}, \eta_0} \cap S_{\kappa_1, \dots, \kappa_d}^{-(n+h_0-1)} \mathfrak{C}_{\mathbf{b}, \eta_0}) \\ & \geq \delta_0^d \sum_{|u_1|, \dots, |u_d| \leq L_0 - 2} \prod_{i=1}^d (L_0 - |u_i| - \mathbf{1}_{u_i \geq 0}) \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \prod_{i=1}^d \psi((n + h_0 - 1) \kappa_i - \Delta_i) \\ & \geq \delta_0^d \cdot L_0^d (L_0 - 1)^d \cdot \left(1 - \frac{1}{L_0} \right)^{2d} \cdot \delta_0^d \cdot \frac{\mathcal{J}_0 N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \\ & = \left(1 - \frac{1}{L_0} \right)^{3d} \cdot \eta_0^{2d} \cdot \frac{\mathcal{J}_0 N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}}. \end{aligned}$$

□

Recall that $E_i = (\gamma_i, \mathfrak{C}_{\mathbf{a}_i, \eta_0})$ where $\gamma_i \in G$. For any $\gamma \in G$, let $V_\gamma = \{1 \leq i \leq t : \gamma_i = \gamma\}$. Since W_0 is a multiple of $|G|$, according to the assumptions (I)-(IV) in

Section 2, $n + h_0 - 1, n + h_1 - 1, \dots, n + h_k - 1$ are all divisible by $|G|$ whenever $n \equiv b \pmod{W}$. So for any $1 \leq i, j \leq t$,

$$\nu(E_i \cap S_\alpha^{-(n+h_0-1)} E_j) = \frac{1}{|G|} \cdot \mathbf{m}_{\mathbb{T}^d}(\mathfrak{C}_{\mathbf{a}_i, \eta_0} \cap S_{\kappa_1, \dots, \kappa_d}^{-(n+h_0-1)} \mathfrak{C}_{\mathbf{a}_j, \eta_0})$$

if and only if i, j belong to the same V_γ . We have

$$\begin{aligned} & \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \cdot \langle g, S_\alpha^{n+h_0-1} g \rangle \\ &= \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \cdot \sum_{1 \leq i, j \leq t} \beta_i \beta_j \cdot \nu(E_i \cap S_\alpha^{-(n+h_0-1)} E_j) \\ &= \sum_{\gamma \in G} \frac{1}{|G|} \sum_{i, j \in V_\gamma} \beta_i \beta_j \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \cdot \mathbf{m}_{\mathbb{T}^d}(\mathfrak{C}_{\mathbf{a}_i, \eta_0} \cap S_{\kappa_1, \dots, \kappa_d}^{-(n+h_0-1)} \mathfrak{C}_{\mathbf{a}_j, \eta_0}) \\ &\geq \left(1 - \frac{1}{L_0}\right)^{3d} \cdot \frac{\mathcal{J}_0 N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \cdot \frac{\eta_0^{2d}}{|G|} \sum_{\gamma \in G} \sum_{i, j \in V_\gamma} \beta_i \beta_j. \end{aligned}$$

By the Cauchy-Schwarz inequality,

$$\begin{aligned} \|g\|_1^2 &= \left(\sum_{\gamma \in G} \frac{\eta_0^d}{|G|} \sum_{i \in V_\gamma} \beta_i \right)^2 \\ &\leq \left(\sum_{\gamma \in G} \frac{1}{|G|} \right) \cdot \left(\sum_{\gamma \in G} \frac{\eta_0^{2d}}{|G|} \left(\sum_{i \in V_\gamma} \beta_i \right)^2 \right) = \frac{\eta_0^{2d}}{|G|} \sum_{\gamma \in H} \sum_{i, j \in V_\gamma} \beta_i \beta_j. \end{aligned}$$

And since $\|f_*\|_1 = \|g_*\|_1$, we have

$$|\|f\|_1 - \|g\|_1| \leq \|f - f_*\|_1 + \|g_* - g\|_1 \leq (2\epsilon_1 + \epsilon_1^2) \|f\|_1,$$

i.e.,

$$\|g\|_1^2 \geq (1 - 2\epsilon_1 - \epsilon_1^2)^2 \|f\|_1^2 \geq \|f\|_1^2 - 4\epsilon_1.$$

When w is sufficiently large, we have

$$\begin{aligned}
 & \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \cdot \langle f, T^{n+h_0-1} f \rangle \\
 & \geq \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \cdot \langle g, S_\alpha^{n+h_0-1} g \rangle - 6\epsilon_1 \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_0) \Omega_n \\
 & \geq \left(\|g\|_1^2 \cdot \left(1 - \frac{1}{L_0}\right)^{3d} - 7\epsilon_1 \right) \cdot \frac{\mathcal{J}_0 N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \\
 & \geq (\|f\|_1^2 - 12\epsilon_1) \cdot \frac{\mathcal{J}_0 N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}}.
 \end{aligned}$$

5. THE PROOF OF THEOREM 1.1

We shall complete the proof of Theorem 1.1.

Proposition 5.1. *Suppose that $A \subseteq \mathcal{B}_X$ with $\mu(A) > 0$ and $0 < \epsilon < \mu(A)^2$. There exist integers $W, b > 0$ with $(b, W) = 1$ and an admissible set $\{h_0, \dots, h_k\}$ such that*

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) \Omega_n \cdot \mu(A \cap T^{-(n+h_i-1)} A) \geq (\mu(A)^2 - \epsilon) \cdot \frac{\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \quad (5.1)$$

for any sufficiently large N and each $0 \leq i \leq k$.

First, assume that T is ergodic. Suppose that $(\mathcal{G}, \mathcal{B}_G, \nu, S_\alpha)$ is the Kronecker factor of $(X, \mathcal{B}_X, \mu, T)$. Write $\mathcal{G} = G_\mu \oplus \mathbb{T}^d$ where G_μ is a finite abelian group.

Suppose that $\{h_0, \dots, h_k\}$ is an admissible set with h_0, \dots, h_k are all the multiple of $|G_\mu|$. Also, assume that W_0 is a positive integer divisible by $|G_\mu|$.

Lemma 5.1. *Suppose that $f \in L_\mu^2$ is a non-negative function. Then for any $\epsilon > 0$, there exist $w_\mu(\epsilon) > 0$ such that for any $w \geq w_\mu(\epsilon)$ and $0 \leq i \leq k$,*

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) \Omega_n \cdot \langle f, T^{n+h_i-1} f \rangle \geq (\|f\|_1^2 - \epsilon) \cdot \frac{\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \quad (5.2)$$

provided that N is sufficiently large, where $W = W_0 \prod_{p \leq w} p$.

Proof. This is the immediate consequence of Propositions 3.1 and 4.1, by noting that $\|\mathbb{E}(f|\mathcal{K})\|_1 = \|f\|_1$. \square

Assume that T is not necessarily ergodic. According to the ergodic decomposition theorem (cf. [7, Theorem 6.2]), there exists a probability space $(\mathcal{Y}, \mathcal{B}_Y, \nu)$ such that

$$\mu = \int_{\mathcal{Y}} \mu_y d\nu(y),$$

and μ_y is a T -invariant ergodic measure on \mathcal{X} for almost every y . Let \mathcal{Y}_0 denote the set of all $y \in \mathcal{Y}$ such that T is ergodic with respect to μ_y .

Let $\epsilon_1 = \epsilon/5$. For any given positive integers w_0 and g_0 , let

$$\mathcal{U}_{w_0, g_0} = \{y \in \mathcal{Y}_0 : w_{\mu_y}(\epsilon_1) = w_0, |G_{\nu_y}| = g_0\}.$$

Note that

$$1 = \mu(\mathcal{X}) = \int_{\mathcal{Y}} \mu_y(\mathcal{X}) dv(y) = \sum_{s, t \geq 1} \int_{\mathcal{U}(s, t)} 1 dv(y).$$

So we may choose sufficiently large w_0 and g_0 such that

$$\sum_{s=1}^{w_0} \sum_{t=1}^{g_0} \int_{\mathcal{U}(s, t)} 1 dv(y) \geq 1 - \epsilon_1.$$

Let

$$\mathcal{Y}_1 = \bigcup_{\substack{1 \leq s \leq w_0 \\ 1 \leq t \leq g_0}} \mathcal{U}_{s, t}.$$

Clearly $v(\mathcal{Y} \setminus \mathcal{Y}_1) \leq \epsilon_1$.

Let $W_0 = [1, 2, \dots, g_0]$ and

$$W = W_0 \prod_{p \leq w_0} p.$$

Choose b, h_0, \dots, h_k satisfying the assumptions (I)-(IV) in Section 2. Furthermore, we may assume that w_0 is sufficiently large such that

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) \Omega_n \leq \frac{2\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}}$$

According to Lemma 5.1, we know that for any $y \in \mathcal{Y}_1$, there exists $N_{\mu_y} > 0$ such that

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) \Omega_n \cdot \mu_y(A \cap T^{-(n+h_i-1)} A) \geq (\mu_y(A)^2 - \epsilon_1) \cdot \frac{\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \quad (5.3)$$

for any $N \geq N_{\mu_y}$. For any positive N_0 , let

$$\mathcal{V}(N_0) = \{y \in \mathcal{Y}_1 : N_{\mu_y} = N_0\}.$$

Then we may choose a large N_0 such that

$$\sum_{r > N_0} \int_{\mathcal{V}(r)} 1 dv(y) \leq \epsilon_1.$$

Let

$$\mathcal{Y}_2 = \bigcup_{1 \leq r \leq N_0} \mathcal{V}(r).$$

Then for any $A_* \in \mathcal{B}_X$, we have

$$\mu(A_*) - \int_{\mathcal{Y}_2} \mu_y(A_*) dv(y) = \int_{\mathcal{Y} \setminus \mathcal{Y}_2} \mu_y(A_*) dv(y) \leq v(\mathcal{Y} \setminus \mathcal{Y}_1) + v(\mathcal{Y}_1 \setminus \mathcal{Y}_2) \leq 2\epsilon_1.$$

Thus (5.3) is valid for any $y \in \mathcal{Y}_2$ and $N > N_0$, So for any $N > N_0$

$$\begin{aligned} & \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) \Omega_n \int_{\mathcal{Y}_2} \mu_y(A \cap T^{-(n+h_i-1)} A) dv(y) \\ & \geq \frac{\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \cdot \left(\int_{\mathcal{Y}_2} \mu_y(A)^2 dv(y) - \epsilon_1 \right). \end{aligned}$$

By the Cauchy-Schwarz inequality,

$$\begin{aligned} \int_{\mathcal{Y}_2} \mu_y(A)^2 dv(y) & \geq \frac{1}{v(\mathcal{Y}_2)} \left(\int_{\mathcal{Y}_2} \mu_y(A) dv(y) \right)^2 \\ & \geq \frac{1}{v(\mathcal{Y}_2)} \left(\int_{\mathcal{Y}} \mu_y(A) dv(y) - 2\epsilon_1 \right)^2 \geq \frac{\mu(A)^2 - 4\epsilon_1}{v(\mathcal{Y}_2)}. \end{aligned}$$

Hence

$$\begin{aligned} & \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) \Omega_n \mu(A \cap T^{-(n+h_i-1)} A) \\ & \geq \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) \Omega_n \int_{\mathcal{Y}_2} \mu_y(A \cap T^{-(n+h_i-1)} A) dv(y) \\ & \geq \frac{\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \left(\frac{\mu(A)^2 - 4\epsilon_1}{v(\mathcal{Y}_2)} dv(y) - \epsilon_1 \right) \\ & \geq (\mu(A)^2 - 5\epsilon_1) \cdot \frac{\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}}. \end{aligned}$$

Proposition 5.1 is concluded. \square

By Proposition 5.1, we may choose $w_0, W_0, b, h_0, \dots, h_k$ such that

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) \Omega_n \cdot \mu(A \cap T^{-(n+h_i-1)} A) \geq \left(\mu(A)^2 - \frac{\epsilon}{4} \right) \cdot \frac{\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}}$$

for any sufficiently large N and each $0 \leq i \leq k$, where $W = W_0 \prod_{p \leq w_0} p$. Furthermore, we may assume that

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \Omega_n \leq \frac{2\mathcal{J}_* N}{(\log R)^{k+1}} \cdot \frac{W^k}{\phi(W)^{k+1}}.$$

and

$$\sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) \Omega_n \leq \left(1 + \frac{\epsilon}{4}\right) \cdot \frac{\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}}.$$

for each $0 \leq i \leq k$. Then

$$\begin{aligned} & \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \varpi(n + h_i) \Omega_n \cdot (\mu(A \cap T^{-(n+h_i-1)} A) - (\mu(A)^2 - \epsilon)) \\ & \geq \left(\mu(A)^2 - \frac{\epsilon}{4}\right) \cdot \frac{\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} - (\mu(A)^2 - \epsilon) \left(1 + \frac{\epsilon}{4}\right) \cdot \frac{\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} \\ & \geq \frac{\epsilon}{2} \cdot \frac{\mathcal{J}_i N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}}. \end{aligned}$$

According to the discussions in [18] and [22], for any $m \geq 1$, we may choose a sufficiently large k and construct a symmetric smooth function $F(t_0, \dots, t_k)$ such that

$$\mathcal{J}_0(F) = \dots = \mathcal{J}_k(F) \geq \frac{10^4 m}{k\epsilon} \cdot \mathcal{J}_*(F).$$

Thus

$$\begin{aligned} & \sum_{\substack{n \sim N \\ n \equiv b \pmod{W}}} \Omega_n \left(\sum_{i=0}^k \varpi(n + h_i) (\mu(A \cap T^{-(n+h_i-1)} A) - (\mu(A)^2 - \epsilon)) - m \log 3N \right) \\ & \geq k \cdot \frac{\epsilon}{2} \cdot \frac{\mathcal{J}_0 N}{(\log R)^k} \cdot \frac{W^k}{\phi(W)^{k+1}} - \frac{2\mathcal{J}_* N \log N}{(\log R)^{k+1}} \cdot \frac{W^k}{\phi(W)^{k+1}} > 0. \end{aligned}$$

Therefore there exists $N \leq n \leq 2N$ such that

$$\sum_{i=0}^k \varpi(n + h_i) (\mu(A \cap T^{-(n+h_i-1)} A) - (\mu(A)^2 - \epsilon)) - m \log 3N > 0.$$

That is, there exist $0 \leq i_0 < i_1 < \dots < i_m \leq k$ such that $n + h_{i_0}, n + h_{i_1}, \dots, n + h_{i_m}$ are all primes and

$$\mu(A \cap T^{-(n+h_{i_j}-1)} A) > \mu(A)^2 - \epsilon$$

for each $0 \leq j \leq m$. Evidently the gaps between those primes are bounded by $\max_{0 \leq i < j \leq k} |h_j - h_i|$.

Remark 5.1. In fact, we may require that the primes p_0, p_2, \dots, p_m in Theorem 1.1 are consecutive primes. Assume that $0 \leq h_0 < h_2 < \dots < h_k$ and write

$$\{a_1, a_2, \dots, a_{h_k-h_0-k}\} = \{h_0, h_0 + 1, \dots, h_k - 1, h_k\} \setminus \{h_0, h_1, \dots, h_k\}.$$

Arbitrarily choose distinct primes $h_k < \rho_1, \rho_2, \dots, \rho_{h_k-h_0-k} \leq w$, which doesn't divide W_0 . Let b satisfy an additional requirement:

$$b \equiv -a_j \pmod{\rho_j}$$

for each $1 \leq j \leq h_k - h_0 - k$. Evidently the above assumption would not contradict with (2.2), since $\rho_j > |h_i - a_j|$ for any $0 \leq i \leq k$. Thus if $n \sim N$ and $n \equiv b \pmod{W}$, for each $1 \leq j \leq h_k - h_0 - k$, $n + a_j$ can not be a prime since it is divisible by ρ_j . It follows that those primes among $n + h_0, n + h_1, \dots, n + h_k$ are surely consecutive primes.

Remark 5.2. In [2], Bergelson, Host and Kra considered the generalizations of Khintchine's theorem for multiple recurrence problems. Suppose that $(\mathcal{X}, \mathcal{B}_\mathcal{X}, \mu, T)$ is a measure-preserving probability system and T is an invertible ergodic transformation. They proved that for any $A \in \mathcal{B}_\mathcal{X}$ with $\mu(A) > 0$ and $\epsilon > 0$, the sets of recurrence

$$\{n \in \mathbb{N} : \mu(A \cap T^{-n}A \cap T^{-2n}A) \geq \mu(A)^3 - \epsilon\}$$

and

$$\{n \in \mathbb{N} : \mu(A \cap T^{-n}A \cap T^{-2n}A \cap T^{-3n}A) \geq \mu(A)^4 - \epsilon\}$$

both have the bounded gaps. However, in general,

$$\{n \in \mathbb{N} : \mu(A \cap T^{-n}A \cap T^{-2n}A \cap T^{-3n}A \cap T^{-4n}A) \geq \mu(A)^5 - \epsilon\}$$

doesn't have a bounded gap.

Let

$$\Lambda_{A,\epsilon}^{(2)} = \{p \text{ prime} : \mu(A \cap T^{-(p-1)}A \cap T^{-2(p-1)}A) \geq \mu(A)^3 - \epsilon\}$$

and

$$\Lambda_{A,\epsilon}^{(3)} = \{p \text{ prime} : \mu(A \cap T^{-(p-1)}A \cap T^{-2(p-1)}A \cap T^{-3(p-1)}A) \geq \mu(A)^4 - \epsilon\}.$$

Naturally, we may ask whether the Maynard-Tao theorem can be extended to $\Lambda_{A,\epsilon}^{(2)}$ and $\Lambda_{A,\epsilon}^{(3)}$. Unfortunately, the exponential sums and the Kronecker factors are not sufficient to attack this problem. According to the work of Host and Kra [14], in order to deal with the multiple recurrence problems, we need to use the characters on the nilmanifolds (cf. [12]) and the factors arising from the Host-Kra seminorms. However, the main difficulty is how to establish a suitable mean-value type summation formula and combine this formula with the Maynard sieve method.

REFERENCES

- [1] A. Balog and A. Perelli, *Exponential sums over primes in an arithmetic progression*, Proc. Amer. Math. Soc., **93**(1985), 578-582.
- [2] V. Bergelson, B. Host and B. Kra, *Multiple recurrence and nilsequences, with an Appendix by I. Ruzsa*, Invent. Math., **160**(2005) 261303.
- [3] V. Bergelson and E. Lesigne, *Van der Corput sets in \mathbb{Z}^d* , Colloq. Math., **110**(2008), 1-49.
- [4] Q. Chu, *Multiple recurrence for two commuting transformations*, Ergodic Theory Dynam. Systems, **31**(2011), 771-792.
- [5] Q. Chu, N. Frantzikinakis and B. Host, *Ergodic averages of commuting transformations with distinct degree polynomials iterates*, Proc. London Math. Soc., **102**(2011), 801-842.
- [6] L. Chua, S. Park and G. D. Smith, *Bounded gaps between primes in special sequences*, Proc. Amer. Math. Soc., **143**(2015), 4597-4611.

- [7] M. Einsiedler, Manfred and T. Ward, *Ergodic theory with a view towards number theory*, Graduate Texts in Mathematics, 259, Springer-Verlag London, Ltd., London, 2011.
- [8] N. Frantzikinakis, *Multiple ergodic averages for three polynomials and applications*, Trans. Amer. Math. Soc., **360**(2008), 5435-5475.
- [9] H. Furstenberg, *Poincaré recurrence and number theory*, Bull. Amer. Math. Soc. (N.S.), **5**(1981), 211-234.
- [10] D. A. Goldston, J. Pintz and C. Y. Yıldırım, *Primes in tuples. I*, Ann. of Math. (2), **170**(2009), 819-862.
- [11] A. Granville, *Primes in intervals of bounded length*, Bull. Amer. Math. Soc. (N.S.), **52**(2015), 171-222.
- [12] B. Green and T. Tao, *The Möbius function is strongly orthogonal to nilsequences*, Ann. of Math. (2), **175**(2012), 541-566.
- [13] H. Helson, *Harmonic analysis*, Addison-Wesley Publishing Company, Advanced Book Program, Reading, MA, 1983.
- [14] B. Host and B. Kra, *Nonconventional ergodic averages and nilmanifolds*, Ann. of Math. (2), **161**(2005), 397-488.
- [15] A. Y. Khintchine, *Eine Verschärfung des Poincaréschen “Wiederkehr-Satzes”*, Compos. Math., **1**(1934), 177-179.
- [16] H.-Z. Li and H. Pan, *Bounded gaps between primes of a special form*, Int. Math. Res. Not., 2015. 12345-12365.
- [17] J. Y. Liu and T. Zhan, *The ternary Goldbach problem in arithmetic progressions*, Acta Arith., **82**(1997), 197-227.
- [18] J. Maynard, *Small gaps between primes*, Ann. of Math. (2), **181**(2015), 383-413.
- [19] P. Pollack, *Bounded gaps between primes with a given primitive root*, Algebra Number Theory, **8**(2014), 1769-1786.
- [20] D. H. J. Polymath, *Variants of the Selberg sieve, and bounded intervals containing many primes*, Res. Math. Sci., **1**(2014), Art. 12, 83 pp.
- [21] A. Sárközy, *On difference sets of sequences on integers III*, Acta Math. Acad. Sci. Hungar., **31**(1978), 355-386.
- [22] T. Tao, *Polymath8b: Bounded intervals with many primes, after Maynard*, <http://terrytao.wordpress.com/2013/11/19/polymath8b-bounded-intervals-with-many-primes-after-maynard>.
- [23] Vinogradov, I. M.: *The method of trigonometrical sums in the theory of numbers*, translated, revised and annotated by K. F. Roth and Anne Davenport, Interscience Publishers, London and New York, 1954
- [24] Y.-T. Zhang, *Bounded gaps between primes*, Ann. of Math. (2), **179**(2014), 1121-1174.

DEPARTMENT OF MATHEMATICS, NANJING UNIVERSITY, NANJING 210093, PEOPLE'S REPUBLIC OF CHINA

E-mail address: haopan79@zoho.com